

DOI: 10.19650/j.cnki.cjsi.J2312176

有限训练样本下的多尺度空洞密集网络 高光谱影像分类*

涂潮¹, 刘万军², 赵琳琳², 曲海成²

(1. 辽宁工程技术大学测绘与地理科学学院 阜新 123000; 2. 辽宁工程技术大学软件学院 葫芦岛 125105)

摘要:为了在有限训练样本情况下充分提取高光谱影像的空间光谱特征,提高分类精度,提出一种结合空洞卷积和密集网络的高光谱影像分类方法。首先,构建多尺度空洞特征提取模块,引入不同数量的空洞卷积层和普通卷积层通过级联的方式增大模型的感受野,并提取多尺度特征。然后,在多尺度空洞特征提取模块之间建立密集连接,实现特征复用的同时缓解梯度消失问题,而模块内部无密集连接,避免构建深度网络而导致网络参数过多的问题。最后,将得到的特征依次通过池化层,全连接层和 Softmax 层完成分类。另外,本文在全连接层后加入 dropout 正则化防止出现过拟合。在 Indian Pines 和 WHU-Hi-Longkou 数据集上与经典分类方法进行对比,本文方法 OA 分别为 98.75% 和 98.82%。实验结果表明,本文设计的网络模型在有限训练样本情况下,分类效果最优。

关键词: 高光谱影像;多尺度特征融合;空洞卷积;密集网络

中图分类号: TH761 文献标识码: A 国家标准学科分类代码: 420.2040

Multiscale dilated dense network for hyperspectral image classification with limited training samples

Tu Chao¹, Liu Wanjun², Zhao Linlin², Qu Haicheng²

(1. School of Geomatics, Liaoning Technical University, Fuxin 123000, China;

2. School of Software, Liaoning Technical University, Huludao 125105, China)

Abstract: In order to fully extract the spatial-spectral features of hyperspectral image with limited training samples and improve classification accuracy, a hyperspectral image classification method combining dilated convolution and dense network is proposed. Firstly, a multi-scale dilated feature extraction module is constructed by introducing different numbers of dilated convolutional layers and ordinary convolutional layers to increase the receptive field of model through cascading and extract multi-scale features. Then, the dense connections are established between multi-scale dilated feature extraction modules to achieve feature reuse while alleviating the problem of gradient vanishing. However, there are no dense connections within the modules to avoid the problem of building a deep network with excessive network parameters. Finally, the obtained features are sequentially classified through pooling layers, fully connected layers, and Softmax layers. In addition, this study adds the dropout regularization after the fully connected layer to prevent overfitting. Compared with classical classification methods on the Indian Pines and WHU-Hi-Longkou datasets, our method provides an OA of 98.75% and 98.82%, respectively. The experimental results show that the network model designed in this study provides the best classification performance at the limited sample conditions.

Keywords: hyperspectral image; multi-scale feature fusion; dilated convolution; dense network

0 引言

高光谱成像是光谱探测技术和成像技术相结合的

一种新型光谱成像技术,包含丰富的空间信息和光谱信息,且具有“图谱合一”的特点,被广泛应用于军事监测、精细农业和矿产资源勘探等领域^[1-3]。例如,在军事上可根据精细光谱特征进行目标识别;在农业上

可根据叶片的光谱变化评估农作物的含水量和叶绿素等营养成分。高光谱影像分类作为实现诸多应用的关键步骤,其目的是对每一个像素分配唯一类别的标签。然而,高光谱影像存在光谱不确定性,高阶非线性性和训练样本有限等关键问题,使得分类任务极具挑战^[4-5]。

早期传统方法主要分为基于光谱信息和基于空谱联合特征两种分类方法^[6]。前者仅利用高光谱影像的光谱信息,借助传统的模式识别方法进行分类,如支持向量机(support vector machine, SVM)^[7]、K-近邻^[8]和稀疏表示^[9]等分类方法。对于地物特征分布复杂的高光谱影像,如农作物光谱谱线相近,存在“异物同谱”现象,仅利用光谱信息分类效果较差。研究学者发现高光谱影像相邻波段像元之间存在着空间联系,因此将空间信息融入到高光谱影像分类方法中,相继提出马尔科夫随机场^[10]、Gabor 滤波^[11]和空间滤波^[12]等分类方法。上述方法虽然取得不错的分类效果,但需要人工且只能提取浅层特征,更依赖于人的先验知识,从而影响特征提取表达和学习能力。

卷积神经网络(convolutional neural network, CNN)因具有局部连接和参数共享等特点,在图像处理领域取得显著效果并使其成功应用到高光谱影像分类^[13]。Hu 等^[14]通过一维卷积层、最大池化层和全连接层搭建简单网络用于高光谱影像分类,与传统方法相比,分类性能得到改善,并为 CNN 用于高光谱影像分类提供新思路。Lyu 等^[15]提出 CNN 与双边滤波结合的分类方法,采用主成分分析法对原始影像进行降维,取第一主成分对输入影像进行滤波,增强空间特征的表达,并将其输入到 CNN 网络进行分类,有效增强模型对边缘区域的分类性能,但降维操作会丢失光谱信息,仅利用空间信息对影像进行分类,效果较差。高光谱影像由一维光谱数据和二维空间数据组成,整体呈三维数据,单独使用光谱信息或空间信息均会影响最终的分类结果。针对这一问题, Li 等^[16]提出三维卷积神经网络(three-dimension convolutional neural network, 3DCNN)用于高光谱影像分类,三维卷积核(three-dimension, 3D)可以在空间维和光谱维均进行卷积操作,同时提取高光谱影像的空间-光谱特征,与同时期的分类方法相比,该方法分类精度最高。研究人员通过设计更深的网络层数来提高分类精度,但随着网络层数的增加,基于 3DCNN 分类方法的网络参数也呈指数级增长。针对这一问题,王燕等^[17]提出快速 3D-CNN 结合深度可分离卷积对高光谱影像进行分类,采用 3DCNN 提取空谱特征,引入深度可分离卷积提取更深层次的空间特征,丰富空谱特征的同时减少了参数量,提高了模型的运行效率,但该方法为了提取更深层次的特征设计深层

网络结构容易出现梯度消失、网络退化的问题。为了解决该问题, Li 等^[18]提出三维残差注意网络模型,以三维卷积神经网络作为基础结构,引入跳跃连接设计残差模块保持参数量不变的情况下提升网络性能,同时嵌入即插即用的注意力机制模块强化重要特征,抑制不重要特征,分类效果较好。为了提高特征的重利用, Li 等^[19]利用三维密集连接网络(three-dimensional-squeeze-and-excitation-densenet, 3D-DenseNet)进行高光谱影像分类,利用密集网络进行深度特征提取,增强相邻卷积层之间特征的重利用,缓解梯度消失问题,并引入 SE 注意力对提取的特征进行重标定,分类性能得到进一步提升。Wang 等^[20]基于密集网络提出快速密集空-谱卷积网络模型(fast dense spectral-spatial convolution network, FDSSC),通过构建 1DCNN 密集块和 3DCNN 密集块串联成深层网络,结合密集网络的优势,取得较好的分类结果。感受野对提升网络分类性能非常重要,为了增大模型的感受野,颜铭靖等^[21]将 3DCNN 与空洞卷积结合提取高光谱影像的空谱特征(three-dimensional dilated convolutional residual neural network, Dilated-3DCNN),但该方法采用单一大小的卷积核无法有效提取高光谱影像的多尺度特征,容易造成单一尺度特征的不完整表达。为了解决该问题,张祥东等^[22]提出基于扩张卷积注意力神经网络的高光谱影像分类方法(dilated convolutional attention neural network, 3D-2D-ADCNN),采用三维-二维卷积神经网络作为基础网络结构,将扩张卷积和注意力结合从而实现不同尺度下特征的融合,以及对重要特征的关注,避免冗余背景信息的干扰,分类性能得到进一步提升。

虽然上述方法在提升高光谱影像分类性能上取得一定的成效,但基于 CNN 的分类方法仍然存在局限性:1)如何在有限训练样本情况下充分提取影像的空谱特征;2)如何在不增加参数的情况下,增大模型的感受野,进一步提取影像的全局特征。针对这两个问题,本文提出一种有限训练样本下的多尺度空洞密集网络(multiscale dilated dense network, MDDN)高光谱影像分类方法。首先,基于三维卷积神经网络设计多尺度空洞特征提取模块,将空洞卷积和普通卷积通过串联的线性方式增大模型的感受野,再以并行的方式将同一层级获取不同感受野的特征进行融合以获取多尺度特征。然后,在模块与模块之间引入密集连接构建深度密集网络,实现相邻模块间的特征复用,并缓解梯度消失问题。最后,将得到的特征依次通过池化层,全连接层和 dropout 层,采用 Softmax 分类器进行分类。在数据集 Indian Pines 和 WHU-Hi-Longkou 上进行实验,与当前主流分类方法相比,本文方法分类性能最优。

1 本文网络结构

1.1 MDDN 网络结构

以 Indian Pines 数据集为例,本文设计的总体网络结构如图 1 所示。首先,以原始高光谱影像标记像元及邻域组成的数据立方体 $7\times7\times200$ 作为网络的输入。其次,设计融合空洞卷积的多尺度空洞特征提取模块,利用三维卷积同时提取高光谱影像的空谱特征,并将不同层次

的空谱特征进行融合,卷积填充方式为“same”,使输入输出的特征图保持一致。然后,在设计的模块与模块之间加入跳跃连接,每一个模块的输入均来自前面所有模块的输出,再以相同的方式传递给后续的模块,实现浅层和深层特征的融合并有效缓解梯度消失问题。另外,模块内部无跳跃连接,减少网络参数,加快网络的收敛速度。最后,将得到的特征依次经过池化层,全连接层和 Softmax 完成分类,在全连接层后添加 dropout 正则化防止出现过拟合。

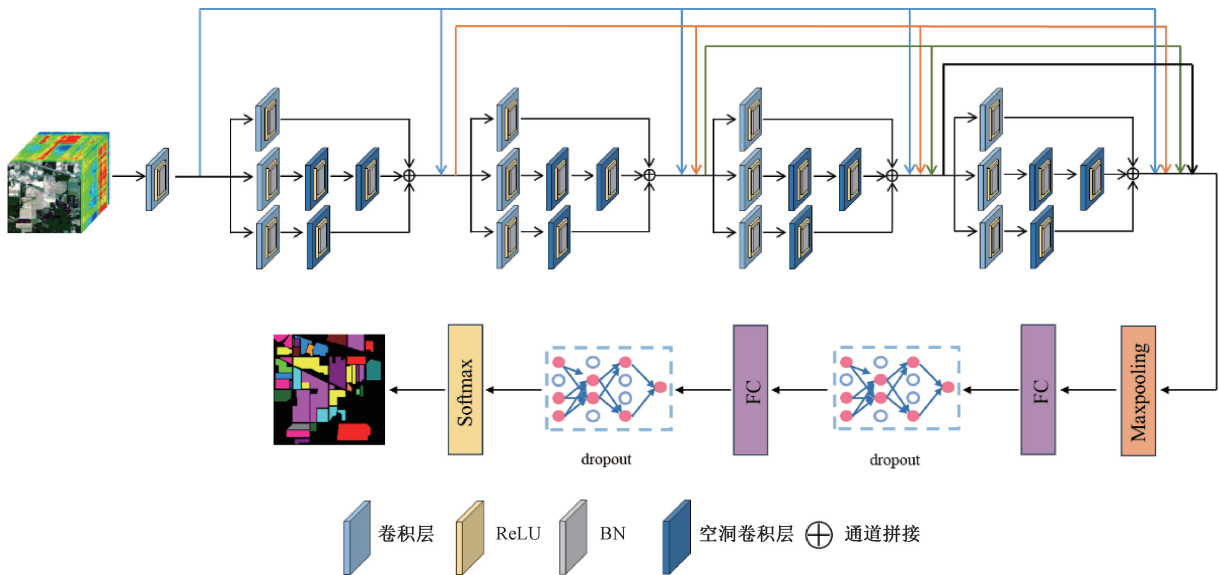


图 1 总体网络结构
Fig. 1 Overall network structure

1.2 多尺度空洞特征提取模块

在图像处理领域,感受野对于提升模型的特征提取能力起到关键作用。通常增大模型感受野的方法有两种:1)增大卷积核的尺寸,但该方法使得网络参数增多,模型计算复杂;2)增加池化层,但池化层通过下采样缩小特征图尺寸来增大模型感受野,使用较多数量的池化层会丢失一部分特征。针对感受野问题,本文引入空洞卷积增大模型感受野的同时不增加网络参数也不损失部分信息。空洞卷积是通过在原始卷积核中相邻两个值之间插入权重为 0 的参数增大卷积核感受野,不同空洞率的卷积核如图 2 所示。

结合空洞卷积和 Inception 网络^[23]构建多尺度空洞特征提取模块 (multiscale dilated dense block, MDD_Block),如图 3 所示。其中,在模块内不同分支设置不同数量的空洞卷积层和普通卷积层串联以获取不同尺度的感受野。

如图 3 所示,对于分支 X_1 :卷积核数量为 32,大小为 $3\times3\times3$,不插入空洞卷积层,根据感受野计算公式 $RF=$

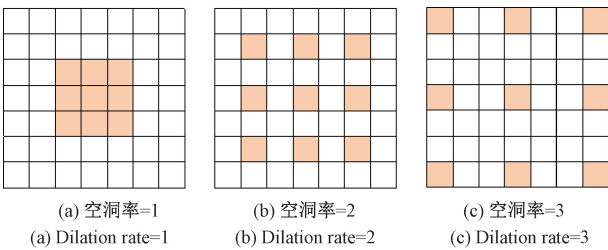


图 2 不同空洞率下的感受野
Fig. 2 The receptive field at different dilation rate

$K+(K-1)(r-1)$, RF 表示感受野, K 表示原始卷积核边长, r 表示空洞率,分支 X_1 感受野大小为 $3\times3\times3$;对于分支 X_2 ,卷积核数量为 32,大小为 $3\times3\times3$,插入空洞卷积层的数量为 2,空洞率为 $2\times2\times2$ (关于空洞率的参数确定见 2.2 节),此时感受野大小为 $11\times11\times11$;对于分支 X_3 ,卷积核数量为 32,大小为 $3\times3\times3$,插入空洞卷积层的数量为 1,空洞率为 $2\times2\times2$,此时感受野大小为 $7\times7\times7$ 。每个分支的输出为:

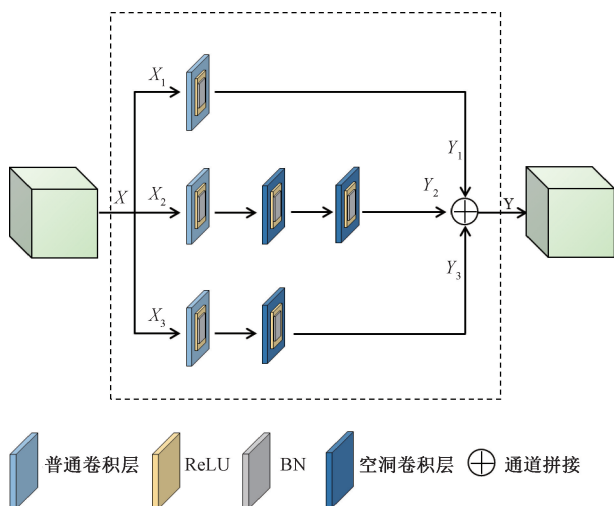


图3 多尺度空洞特征提取模块

Fig. 3 Multiscale dilated feature extraction module

$$Y_i = \begin{cases} C(X_i), & i = 1 \\ D(D(C(X_i))), & i = 2 \\ D(C(X_i)), & i = 3 \end{cases} \quad (1)$$

式中: $C(\cdot)$ 表示卷积操作, $D(\cdot)$ 表示空洞卷积操作, 在卷积层后加入 ReLU 和 BN 层以缓解梯度消失, 加快收敛。模块内共有 3 条分支分别提取不同尺度的特征, 然后通过拼接操作进行特征融合。

1.3 密集网络

高光谱影像光谱维度高, 相邻波段间存在冗余特征, 较深的网络结构容易出现网络退化、梯度消失的问题。为了进一步增强网络的特征提取能力, 本文在构建的 MDD_Block 之间引入密集连接, 设计多尺度密集网络以提取影像的深层语义信息, 每个模块的输入均来自前面所有模块的特征信息, 并以相同的方式传递到后续的模块, 其结构如图 4 所示。

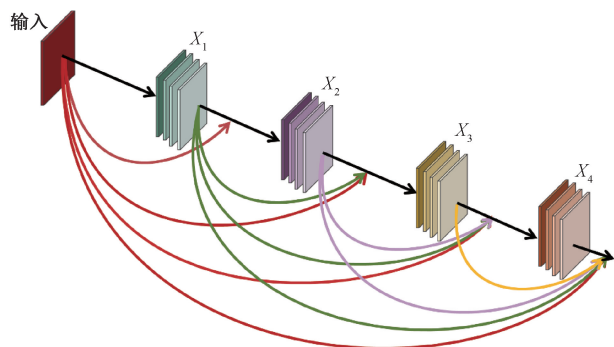


图4 密集网络

Fig. 4 Dense network

从图 4 可以看出, 该网络由 4 个密集连接的多尺度空洞特征提取模块组成, 建立不同模块之间的连接关系,

充分利用每一个模块的特征图, 缓解了梯度消失问题。每一层的连接方式可表达为:

$$M_i = N_i[M_0, M_1, \dots, M_{i-1}] \quad (2)$$

式中: M_i 为第 i 层的输出; $[M_0, M_1, \dots, M_{i-1}]$ 为第 0 层到第 $i-1$ 层的所有特征图的输出; $N_i[\cdot]$ 为 3 个连续操作的复合函数: 批量归一化 (BN)、线性单元 (ReLU)、Conv ($3 \times 3 \times 3$)。

2 实验结果和分析

实验所用的硬件环境为 R9-5900HX、NVIDIA GeForce RTX3060 显卡、16 GB 内存, 在 Windows10 系统下基于 Python 并采用 keras 和 Tensorflow 两种框架构建网络模型。

2.1 数据集

Indian Pines 数据集由传感器 AVRIS 在美国印第安纳州获取的影像, 空间分辨率 20 m, 影像大小为 145 pixel $\times$ 145 pixel, 波段数目为 224 个, 剔除受水汽吸收和低信噪比的波段后, 保留剩余的 200 个波段。该数据集共包含 16 个类别, 有 10 249 个样本数据。每类地物随机选取 50 个样本并将其按照 7:3 的比例分为训练集和验证集 (其中类别 1、类别 7 和类别 9 样本总数不足 50, 则选取该类别的 50% 按照 7:3 的比例分为训练集和验证集), 其余的作为测试集。

WHU-Hi-Longkou 数据集^[24-25]为 RSIDEA 研究团队于 2018 年 7 月 17 日对湖北省龙口镇获取的高光谱影像数据, 影像大小为 550 pixel $\times$ 400 pixel, 波段数目为 270 个, 光谱范围为 400~1 000 nm, 光谱分辨率和空间分辨率分别为 6 nm 和 0.463 m, 有 9 种类别, 包含 204 542 个已标记地物类别的像元。每类地物分别随机选取 0.1% 样本作为训练集, 0.1% 样本作为验证集, 其余的作为测试集。

2.2 网络参数设置

本文采用随机梯度下降法更新权重, 学习率为 0.001, batch_size 为 32, epoch 为 100, 卷积核大小为 $3 \times 3 \times 3$, 步长为 (1, 1, 2), 卷积核数量 32 个, dropout 设置为 0.5。影响本文方法的网络参数主要包括: 多尺度空洞特征提取模块的数量、空间尺寸和空洞率, 本文将针对这些参数进行相应的实验, 依次分析确定最优参数。

1) 多尺度空洞特征提取模块 (MDD_Block): 本文将不同数量的空洞卷积层和普通卷积层串联实现在不增加参数的同时增大模型的感受野, 构建并行分支网络提取不同感受野下的特征, 提升网络分类性能。如果模块的数量过少则提取不到关键特征, 如果模块的数量过多则

容易出现网络退化,导致精确率下降。因此,本文采用 OA 作为评价指标进行实验分析,实验结果如表 1 所示。当 MDD_Block 的数量为 3 时,OA 分别为 96.86%、97.16%,分类效果较差。当 MDD_Block 的数量为 4 时,精确率上升,说明本文设计的特征提取模块对于提升网络性能起到一定的作用。当 MDD_Block 的数量为 5 时,OA 分别为 97.33%、97.85%,说明过深的网络结构出现网络退化的问题,导致精确率下降。综合考虑,当 MDD_Block 的数量为 4 时,效果最优。

表 1 不同数量 MDD_Block 下的分类精度
Table 1 Classification accuracy under different number of MDD_Block

多尺度空洞特征 提取模块数量	OA/%	
	Indian Pines	WHU-Hi-Longkou
3	96.86	97.16
4	97.51	97.92
5	97.33	97.85

2) 空间尺寸(patch_size):空洞卷积在卷积核的元素之间添加间隔,如果 patch 尺寸过小,空洞卷积与普通卷积串联后感受野范围大于 patch 尺寸,则会忽略细节特征的提取。如果 patch 尺寸过大,会带来冗余特征的干扰。因此,选择合适的 patch 尺寸对提升模型的性能非常重要,本文采用 OA 和训练时间两个评价指标进行实验分析,实验结果如表 2 所示。

表 2 不同空间尺寸下的分类精度
Table 2 Classification accuracy at different spatial sizes

空间尺寸	Indian Pines		WHU-Hi-Longkou	
	OA/%	时间/s	OA/%	时间/s
3×3	97.78	948.6	98.01	1 472.5
5×5	97.91	1 170.6	98.30	1 793.2
7×7	98.35	1 533.1	98.42	2 256.7
9×9	98.37	2 056.9	98.47	2 698.3
11×11	98.36	2 641.9	98.46	3 278.6

以 Indian Pines 数据集为例,由表 2 可知,当输入的空间尺寸分别为 3×3 和 5×5 时,精确率较低,说明输入空间尺寸过小时网络学习到的特征也相对较少,从而影响分类结果。当输入的空间尺寸为 7×7 时,精确率出现上升趋势,说明网络对于特征的提取依赖于空间尺寸的大小。当输入的空间尺寸为 9×9 时,精确率的提升控制在 0.03% 以内,训练时间较长。当输入的空间尺寸更大时,带来较多冗余特征的干扰,因此精确率

下降。综上分析,对于 Indian Pines 输入的空间尺寸选取 7×7,WHU-Hi-Longkou 数据集输入的空间尺寸选取 9×9。

3) 空洞率:空洞卷积通过在原始卷积核内行与列之间添加权重为 0 的方式增大模型感受野,但行与列采样间隔过大容易忽略影像的细节特征,而这些细节特征很有可能是判别地物类间差异的关键因素。因此,需要选取合适的空洞率,高光谱影像作为三维数据,本文分别从空间维和光谱维进行实验,采用 OA 和 Kappa 两个评价指标,实验结果如图 5 所示。

如图 5(a) 所示,对于 Indian Pines 数据集,当空间维选择 2 时,总体分类精度最高,这表明空洞卷积可以发挥一定作用,但当空洞率选择较大参数时,总体分类精度曲线呈下降趋势,由于本文输入的空间尺寸为 7×7,此时卷积核的感受野尺寸远大于输入的空间尺寸,导致忽略相邻特征之间的联系,因此分类精度下降;光谱维曲线重复出现上升下降的情况,这是因为空洞率光谱维选择较大参数,此时学习到更多的特征,因此呈现上升趋势,但当空洞参数过大时,此时丢失的特征会多于学习到的特征,因此呈现下降的趋势。综合考虑,本文对于 Indian Pines 数据集,空洞率选择的参数为 2×2×2,对于 WHU-Hi-Longkou 数据集,空洞率选择的参数为 2×2×4。

2.3 对比实验

为了验证本文网络模型设计的有效性,选取 SP-SVM^[7]、3DCNN^[16]、3D-DenseNet^[19]、FDSSC^[20]、Dilated-3DCNN^[21]和 3D-2D-ADCNN^[22]作为对比方法进行实验。为了保证实验的公平性和可参照性,所选取对比方法的参数与原文保持一致,训练样本和测试样本数量与本文保持一致,所有实验重复 10 次,所有方法的最终分类结果均为平均值。各方法在 Indian Pines 数据集上的分类结果如表 3 所示,加粗字体表示结果最优。

由表 3 可知,本文提出的 MDDN 方法共有 9 种地物的分类精度最高,其中有 4 种地物达到完全正确的分类,其余 7 种地物的分类精度也都在 96% 以上。相比于其他 6 种方法,OA 分别提高了 32.1%、5.89%、1.12%、2.71%、2.84%、1.13%。为了更直观表现本文方法的有效性,图 6 所示为不同方法在 IN 数据集上的分类结果。

从图 6 可以看出,SP-SVM 方法分类效果最差,其原因是仅利用光谱信息进行识别分类,忽略高光谱影像的空间信息,导致各类地物均存在大片错误分类。3DCNN、FDSSC、3D-DenseNet、Dilated-3DCNN、3D-2D-ADCNN 和 MDDN 方法由于同时利用影像的空谱特征,地物分类效果相较于 SP-SVM 方法得到明显提升,但 3DCNN 分类精度最低,这是因为在有限训练样本条件下,通过增加网

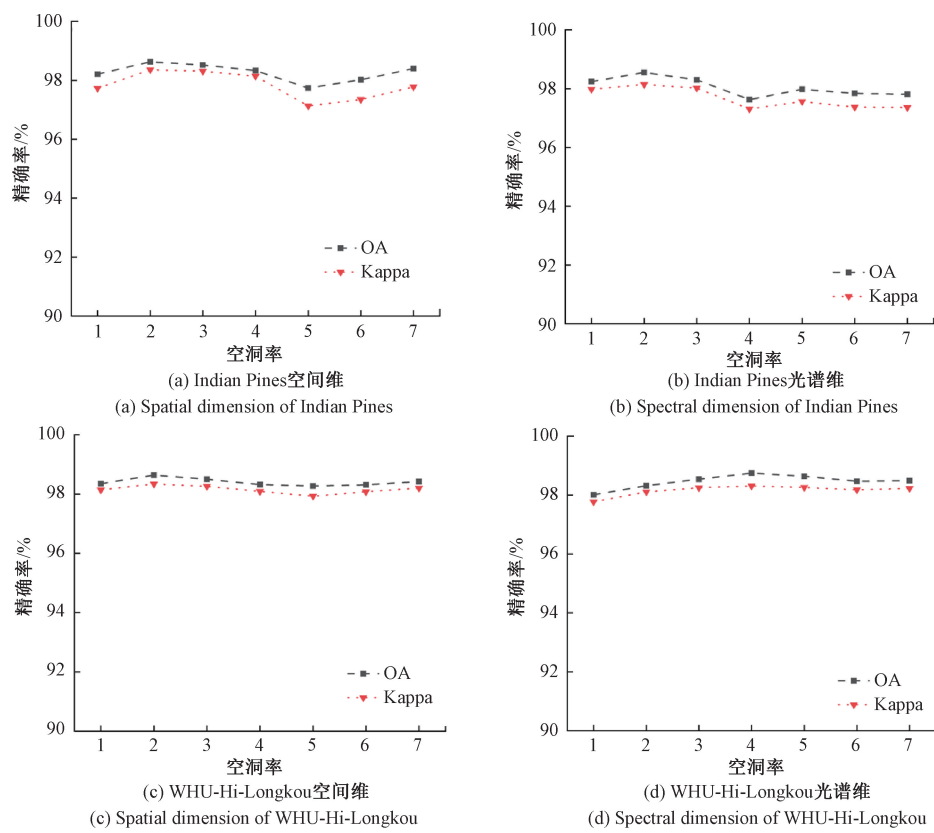


图 5 不同空洞率下的分类精度

Fig. 5 Classification accuracy under different dilation rate

表 3 不同方法在 Indian Pines 数据集上的分类结果

Table 3 Classification results of different methods on the Indian Pines dataset							%
类别	SP-SVM	3DCNN	FDSSC	3D-DenseNet	Dilated-3DCNN	3D-2D-ADCNN	MDDN
1	42.72	100	100	97.67	97.22	99.00	100
2	35.66	96.71	97.80	97.69	98.10	97.12	97.87
3	57.54	91.58	96.55	95.22	94.24	97.05	97.63
4	63.00	84.15	99.58	96.15	96.90	98.32	96.77
5	49.88	99.03	97.30	99.25	98.64	98.72	97.90
6	63.02	97.56	96.94	97.79	97.50	98.89	99.79
7	82.45	46.00	96.15	73.68	88.42	100	99.05
8	77.78	97.94	99.79	99.79	97.13	100	100
9	97.68	100	100	94.74	83.01	94.11	99.79
10	33.90	87.25	97.07	95.76	94.44	96.68	100
11	48.92	95.14	97.34	95.72	95.17	98.02	99.18
12	76.48	86.56	95.45	94.66	92.47	91.89	99.31
13	53.96	97.62	99.03	100	97.56	100	97.15
14	79.69	96.22	99.21	94.82	98.01	99.12	100
15	92.38	73.92	97.10	89.44	88.28	93.61	99.29
16	64.37	98.80	100	100	98.81	100	99.74
OA	66.65	92.86	97.63	96.04	95.91	97.62	98.75
AA	63.71	90.53	98.08	95.15	94.74	97.66	98.97
Kappa	62.94	91.86	97.30	95.48	95.33	97.29	98.57

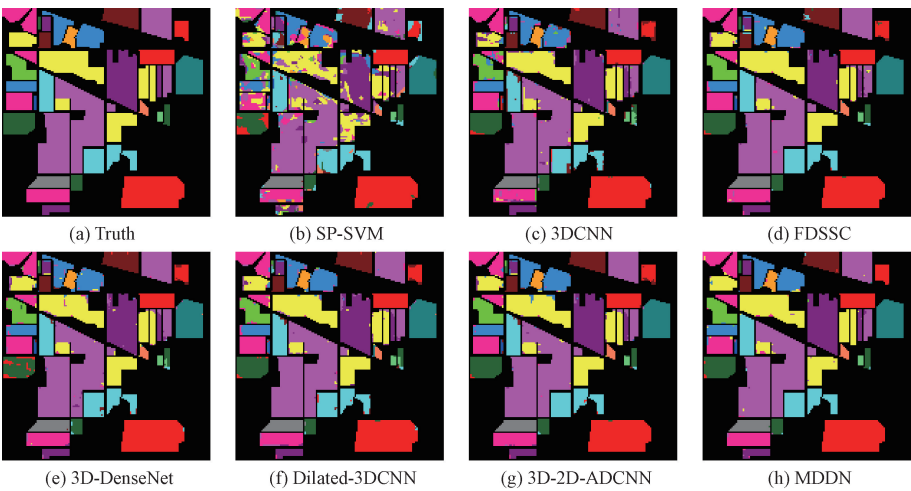


图 6 不同方法在 Indian Pines 数据集上的分类结果

Fig. 6 Classification results of different methods on the Indian Pines dataset

络层数提高分类精度容易出现网络退化的问题。Dilated-3DCNN 方法引入空洞卷积和残差网络,分类精度得到提升,说明空洞卷积和残差网络的结合使用对于提升网络性能能够起到一定的作用,但从图 6 可以看出仍然存在较多的噪点,以及小范围内的大片错分。FDSSC 和 3D-DenseNet 方法均使用密集网络,前者通过 1DCNN 密集块和 3DCNN 密集块串联成深层网络,采用不同大小的卷积核分别提取空间和光谱特征;后者使用三维卷积核同时提取影像的空谱特征,引入密集网络进行更深层特征的提取,并利用通道注意力机制重新校准不同通道间的关系,但两种方法在地物边缘处仍然存在大量噪点。3D-2D-ADCNN 结合空洞卷积和两种不同的注意力机制,

使得分类性能进一步提升,但在有限训练样本条件下无法准确对所有地物实现精确分类,存在较多的孤立噪点,且对于样本总数较少的地物类别分类效果较差,如类别 9、12 和 15。本文提出的 MDDN 方法能够较为准确实现各类地物的正确分类,整体区域相对平滑,仅有较少的孤立噪点。

2.4 实测实验与分析

在本节中我们采用武汉大学公开共享的无人机高光谱高空间(双高)遥感影像数据集 WHU-Hi-Longkou 作为实测数据集,进一步验证本文方法的实用性。各方法在 WHU-Hi-Longkou 数据集的分类结果如表 4 所示,加粗字体表示最优结果。

表 4 不同方法在 WHU-Hi-Longkou 数据集上的分类结果

Table 4 Classification results of different methods on the WHU-Hi-Longkou dataset

类别	SP-SVM	3DCNN	FDSSC	3D-DenseNet	Dilated-3DCNN	3D-2D-ADCNN	MDDN	%
1	53.24	66.93	90.83	85.88	74.39	94.16	96.02	
2	94.78	97.01	95.35	97.84	98.14	98.92	99.29	
3	46.94	48.24	88.77	86.21	62.19	89.57	95.24	
4	70.21	91.59	94.09	88.69	96.13	93.66	97.51	
5	89.21	86.63	96.65	86.90	92.69	98.26	99.03	
6	29.79	79.06	82.74	80.23	80.06	77.85	96.23	
7	91.57	77.47	94.18	99.27	95.16	99.30	99.46	
8	99.94	99.85	98.97	96.98	99.96	99.96	99.95	
9	96.46	97.21	95.99	94.92	93.42	95.89	97.61	
OA	89.17	90.13	97.37	96.54	94.08	97.91	98.82	
AA	74.68	82.67	93.07	90.78	88.00	94.17	97.70	
Kappa	85.71	86.90	96.55	95.46	92.21	97.25	98.45	

从表 4 可以看出,在实测数据集 WHU-Hi-Longkou 上,本文提出的 MDDN 方法仍然达到最高的分类精度,OA、AA 和 Kappa 分别为 98.82%、97.70% 和 98.45%。相比于其他 6 种方法,OA 分别提高了 9.65%、8.69%、1.45%、2.28%、4.74%、0.91%。不同方法在 WHU-Hi-Longkou 数据集上的分类结果如图 7 所示。SP-SVM 方法只能利用高光谱影像的光谱特征,存在大量的椒盐噪点。相较而言,基于 CNN 的分类方法具有更好的分类效果。3DCNN 引入空间信息,同时提取高光谱影像的空谱特征,分类性能优于 SP-SVM,

但在图 7 上仍然存在大片的错误分类。FDSSC 和 3D-DenseNet 引入密集网络,增强了上下层特征的传递,缓解因网络层数较多带来的梯度消失问题,但对于边缘区域的分类效果较差,都存在大量的错误分类现象。Dilated-3DCNN 和 3D-2D-ADCNN 都利用空洞卷积增大模型的感受野,进一步提升了分类性能,但同样在边缘区域存在较多的误分类点。本文提出的 MDDN 方法在整体上孤立噪点较少,在提取复杂空间细节特征方面效果也优于其他 6 种方法,分类结果图更接近真实影像。

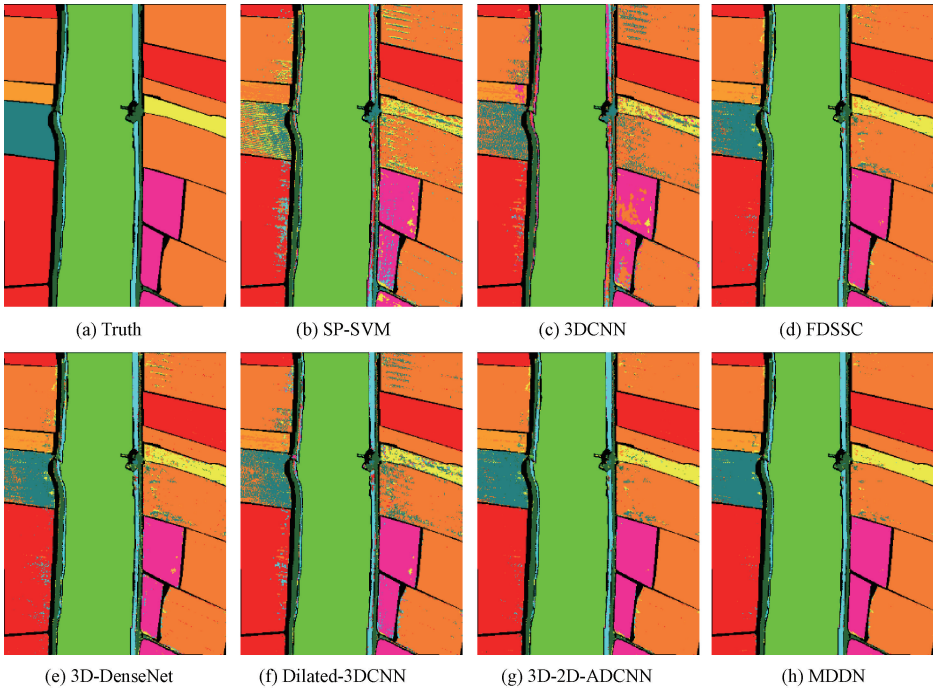


图 7 不同方法在 WHU-Hi-Longkou 数据集上的分类结果

Fig. 7 Classification results of different methods on the WHU-Hi-Longkou dataset

2.5 运行性能实验分析

为了比较本文提出的 MDDN 方法与 SP-SVM、3DCNN、FDSSC、3D-DenseNet、Dilated-3DCNN 和 3D-2D-ADCNN 方法的效率,本文在统一的硬件环境下进行实验测试。硬件环境为 R9-5900HX、NVIDIA GeForce RTX3060 显卡、16 GB 内存。训练集、验证集和测试集样本数量的选取参照 2.1 节,网络参数的设置参照 2.2 节。

表 5 为不同方法在 Indian Pines 和 WHU-Hi-Longkou 两个数据集上的训练时间和测试时间。

从表 5 可以看出,在精度上本文提出的 MDDN 方法精度最高,效果最好。在效率上,与 SP-SVM、3DCNN 方法相比,本文方法处理时间长,这是因为 SP-SVM 方法仅利用光谱特征,提取的特征数量较少;3DCNN 包含的卷积层数量较少,只能提取浅层特征,无法提取更深

表 5 不同方法在两个数据集上的训练时间和测试时间

Table 5 Training and testing time of different methods on two datasets

数据集	时间/s	SP-SVM	3DCNN	FDSSC	3D-DenseNet	Dilated-3DCNN	3D-2D-ADCNN	MDDN
Indian Pines	训练时间	13. 78	156. 32	582. 67	1 982. 62	1 358. 41	1 195. 99	1 533. 14
	测试时间	0. 83	1. 54	6. 45	15. 83	14. 79	12. 59	13. 56
WHU-Hi-Longkou	训练时间	184. 76	583. 65	1 336. 29	3 623. 81	2 489. 03	2 053. 86	2 744. 52
	测试时间	3. 83	6. 68	10. 34	48. 07	21. 51	20. 26	23. 04

层次的特征,运算量较少。与 3D-DenseNet 方法相比,本文方法分类精度高且运行效率更低。与深层网络 FDSSC、Dilated-3DCNN 和 3D-2D-ADCNN 方法相比,本文方法的训练时间较长,这是由网络框架决定的。首先,本文构建的深层密集网络在结构上比其他 3 种方法使用的 CNN 更深,可以充分利用网络中的浅层细节特征和深层语义特征。其次,由多尺度空洞特征提取模块组成的网络可以利用更复杂的空间信息。这两方面保证了本文方法能够提取更具有辨别性的空间特征,并且本文方法的测试时间较短,能够满足实际应用的需求。

3 结 论

针对现有基于 CNN 分类方法在有限训练样本情况下无法充分提取高光谱影像的空谱特征,导致分类精度低的问题,提出了结合空洞卷积和密集网络的高光谱影像分类方法。引入空洞卷积构建多尺度空洞特征提取模块,实现不同层次下空谱特征的融合。由于多个空洞卷积层叠加,采用相同采样间隔容易出现网格效应,本文采用不同数量的空洞卷积层和普通卷积层进行串联缓解网格效应,并针对高光谱影像的高维特性,分别从空间维和光谱维选取最优的空洞率,进一步提取空间光谱特征。在构建的模块之间加入跳跃连接,构建深度密集网络,实现不同模块之间的特征重用,加强特征传播并缓解梯度消失问题。在两个数据集上进行实验,与其他方法相比,结果证明,本文提出的 MDDN 方法在训练样本较少的条件下分类性能最优。

考虑到空洞率的选择需要人工多次实验确定最优参数,下一步将针对模型如何实现自适应感受野进行研究。同时,现有基于 CNN 分类方法构建更宽更深的网络模型提高分类精度导致网络参数增多,如何构建轻量化网络也在下一步研究工作当中。

参考文献

[1] 余旭初,刘冰,薛志祥. 高光谱地物要素识别潜力分析与前景展望[J]. 测绘学报, 2023, 52(7): 1115-1125.

YU X CH, LIU B, XU ZH X. Potential analysis and prospect of hyperspectral ground object recognition[J]. Acta Geodaetica et Cartographica Sinica, 2023, 52(7): 1115-1125.

[2] 龚威,史硕,陈博文,等. 机载高光谱激光雷达成像技术发展与应用[J]. 光学学报, 2022, 42(12): 29-40.

GONG W, SHI SH, CHEN B W, et al. Development and application of airborne hyperspectral LiDAR imaging technolog[J]. Acta Optica Sinica, 2022, 42(12): 29-40.

[3] 苏红军. 高光谱遥感影像降维:进展、挑战与展望[J]. 遥感学报, 2022, 26(8): 1504-1529.

SU H J. Dimensionality reduction for hyperspectral remote sensing advances, challenges, and prospects[J]. National Remote Sensing Bulletin, 2022, 26(8): 1504-1529.

[4] 叶珍,白璘,何明一. 高光谱图像空谱特征提取综述[J]. 中国图象图形学报, 2021, 26(8): 1737-1763.

YE ZH, BAI L, HE M Y. Review of spatial-spectral feature extraction for hyperspectral image[J]. Journal of Image and Graphics, 2021, 26(8): 1737-1763.

[5] 孙根云,付航,张爱竹,等. 高光谱影像奇异谱分析特征提取方法:综述与评价[J]. 测绘学报, 2023, 52(7): 1148-1163.

SUN G Y, FU H, ZHANG AI ZH, et al. Singular spectrum analysis method for hyperspectral imagery feature extraction: A review and evaluation[J]. Acta Geodaetica et Cartographica Sinica, 2023, 52(7): 1148-1163.

[6] 万升,杨健,宫辰. 基于图神经网络的高光谱图像分类研究进展[J]. 电子学报, 2023, 51(6): 1687-1709.

WANG SH, YANG J, GONG CH. Advances of hyperspectral image classification based on graph neural networks[J]. Acta Electronica Sinica, 2023, 51(6): 1687-1709.

[7] 李玉,宫学亮,赵泉华. 基于张量径向基核函数支持向量机的高光谱影像分类[J]. 仪器仪表学报, 2020, 41(12): 253-262.

LI Y, GONG X L, ZHAO Q H. Hyperspectral image classification based on tensor-based radial basis kernel function and support vector machine[J]. Chinese Journal of Scientific Instrument, 2020, 41(12): 253-262.

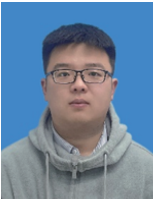
[8] LI D, SHEN Y, KONG F, et al. Spectral-spatial prototype learning-based nearest neighbor classifier for hyperspectral images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-15.

[9] 陈善学,何宇峰. 基于空谱字典的加权联合稀疏表示高光谱图像分类[J]. 光学学报, 2023, 43(1):

- 66-76.
- CHEN SH X, HE Y F. Weighted joint sparse representation hyperspectral image classification based on spatial-spectral dictionary [J]. *Acta Optica Sinica*, 2023, 43(1): 66-76.
- [10] PAN C, JIA X, LI J, et al. Adaptive edge preserving maps in markov random fields for hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(10): 8568-8583.
- [11] 吕欢欢, 胡杨, 张辉. 三维 Gabor 和多尺度残差网络的高光谱影像分类[J]. *遥感信息*, 2023, 38(4): 33-41.
- LYU H H, HU Y, ZHANG H. Hyperspectral image classification based on three-dimensional gabor and multi-scale residual network[J]. *Remote Sensing Information*, 2023, 38(4): 33-41.
- [12] CHEN G Y, KRZYSAK A, QIAN S E. Hyperspectral imagery classification with minimum noise fraction, 2D spatial filtering and SVM [J]. *International Journal of Wavelets, Multiresolution and Information Processing*, 2022, 20(6): 2250025.
- [13] 易琰, 张宇航, 宗艳桃, 等. 基于卷积神经网络的高光谱图像分类算法综述[J]. *电光与控制*, 2023, 30(3): 70-77.
- YI Q, ZHANG Y H, ZONG Y T, et al. A survey of hyperspectral image classification algorithms based on convolutional neural networks[J]. *Electronics Optics & Control*, 2023, 30(3): 70-77.
- [14] HU W, HUANG Y Y, WEI L, et al. Deep convolutional neural networks for hyperspectral image classification[J]. *Journal of Sensors*, 2015, 2015(3): 1-12.
- [15] LYU H H, WANG Z H, ZHANG H. Edge protection filtering and convolutional neural network for hyperspectral remote sensing image classification [J]. *Infrared Physics and Technology*, 2022, 122, 104039.
- [16] LI Y, ZHANG H K, SHEN Q. Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network [J]. *Remote Sensing*, 2017, 9(1): 1-21.
- [17] 王燕, 梁琦. 快速 3D-CNN 结合深度可分离卷积对高光谱图像分类[J]. *计算机科学与探索*, 2022, 16(12): 2860-2869.
- WANG Y, LIANG Q. Fast 3D-CNN combined with depth separable convolution for hyperspectral image classification [J]. *Journal of Frontiers of Computer Science and Technology*, 2022, 16(12): 2860-2869.
- [18] LI H Z, WEI K H, ZHANG B. 3D residual attention network for hyperspectral image classification [J]. *International Journal of Wavelets Multiresolution and Information Processing*, 2023, 21(4): 1.
- [19] LI G D, ZHANG C J, LEI R M, et al. Hyperspectral remote sensing image classification using three-dimensional-squeeze-and-excitation-densenet (3D-SE-DenseNet)[J]. *Remote Sensing Letters*, 2020, 11(2): 195-203.
- [20] WANG W J, DOU S G, JIANG Z M, et al. A fast dense spectral-spatial convolution network framework for hyperspectral image classification [J]. *Remote Sensing*, 2018, 10(7): 1068.
- [21] 颜铭靖, 苏喜友. 基于三维空洞卷积残差神经网络的高光谱影像分类方法[J]. *光学学报*, 2020, 40(16): 1628002.
- YAN M J, SU X Y. Hyperspectral image classification based on three-dimensional dilated convolutional residual neural network [J]. *Acta Optica Sinica*, 2020, 40(16): 1628002.
- [22] 张祥东, 王腾军, 朱劭俊, 等. 基于扩张卷积注意力神经网络的高光谱图像分类[J]. *光学学报*, 2021, 41(3): 49-59.
- ZHANG X D, WANG T J, ZHU SH J, et al. Hyperspectral image classification based on dilated convolutional attention neural network [J]. *Acta Optica Sinica*, 2021, 41(3): 49-59.
- [23] 高红民, 朱敏, 曹雪莹, 等. 多尺度融合注意力机制的胆囊癌显微高光谱图像分类[J]. *中国图象图形学报*, 2023, 28(4): 1173-1185.
- GAO H M, ZHU M, CAO X Y, et al. A micro-hyperspectral image classification method of gallbladder cancer based on multi-scale fusion attention mechanism [J]. *Journal of Image and Graphics*, 2023, 28(4): 1173-1185.
- [24] ZHONG Y F, HU X, LUO C, et al. WHU-Hi: UAV-borne hyperspectral with high spatial resolution (H2) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with CRF [J]. *Remote Sensing of Environment*, 2020, 250: 11201.
- [25] ZHONG Y F, WANG X Y, XU Y, et al. Mini-UAV-

borne hyperspectral remote sensing: from observation and processing to applications [J]. IEEE Geoscience and Remote Sensing Magazine, 2018, 6(4):46-62.

作者简介



涂潮(通信作者),2018 年于无锡太湖学院获得学士学位,2021 年于辽宁工程技术大学获得硕士学位,现为辽宁工程技术大学博士生,主要研究方向高光谱影像分类、图像处理 and 计算机视觉。

E-mail:745700558@qq.com

Tu Chao (Corresponding author) received his B. Sc. degree from Taihu University of Wuxi in 2018, received his M. Sc. degree from Liaoning Technical University in 2021. He is currently a doctoral student at Liaoning Technical University. His

main research interest is hyperspectral image classification, image processing, and computer vision.



刘万军,1982 年于燕山大学获得学士学位,1991 年于辽宁工程技术大学获得硕士学位,现为辽宁工程技术大学教授,博士生导师,主要研究方向计算机视觉、模式识别与人工智能。

E-mail: 2149541390@qq.com

Liu WanJun received his B. Sc. degree from Yanshan University in 1982, received his M. Sc. degree from Liaoning Technical University in 1991. He is currently a professor and a Ph. D. advisor at Liaoning Technical University. His main research interest is computer vision, pattern recognition and artificial intelligence.